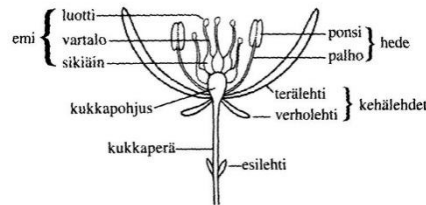
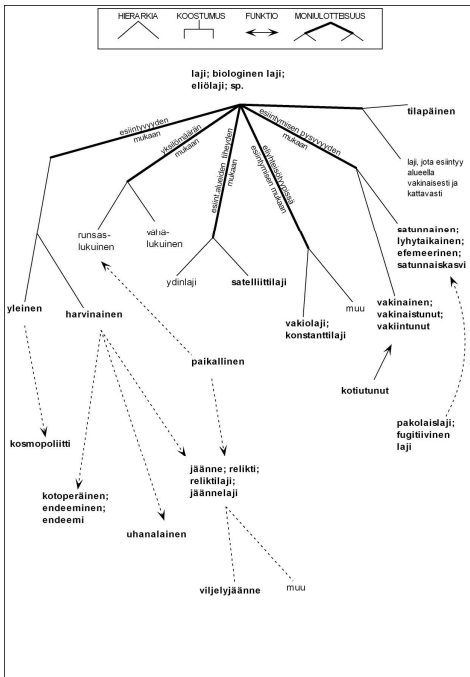


# Possibilities for term descriptions in a wiki-based terminological platform

Kaarina Pitkänen-Heikkilä

# The Bank of Finnish Terminology in Arts and Sciences



machine translation | MT

### Definition (en)

1. sub-field of language technology that investigates the use of software to translate text or speech
2. translation made by **machine translation program**

Categories: [Biodiversity](#) | [Conservation biology](#) | [Population genetics](#) | [Species](#) | [Ecology](#)

## WIKI-BASED PLATFORM

- collaborative platform where edits are showed immediately
- demands from authors courage to edit entries and tolerating incompleteness

## EXPERTS' NICHESOURCING

- experts have the best knowledge from the concepts of their own special field
- experts are responsible for the accurate description of their terms and concepts

## MASSIVE TERMBANK

- How manage that all terminological information?
- How can all the terminological information be managed?

## SUPPORT FOR TERMINOLOGY WORK

- How can we support the terminology work of experts in the best way?
- How can ordinary users participate in crowdsourced terminology work?

# Wiki platform vs. traditional dictionary

## WIKI IS A COLLECTIVE AND TRANSPARENT PLATFORM

- discussion is possible on every page
- authors can be identified from the page history
- challenge: single edits of experts or users can be hidden in the massive data

## SEPARATE NAMESPACES FOR DIFFERENT DOMAINS



- each expert can only participate in her/his own domain's terminology work
- many kind of overlapping can be described
  - concepts can be bound to part of a concept system (or many different systems)
  - the platform brings up homonymy, polysemy and synonymy

## GENERAL BENEFITS OF A DIGITAL FORMAT


- an easy operating system


# Designation page in the BFT


[Designation](#) [Discussion](#)


[Read](#) [Edit](#) [Edit source](#) [View history](#)  [More ▾](#)  

## *anomalía*

**anomalía** (*Filosofia*) *ristiriitainen havainto suhteessa paradigmaan*  
Lue koko artikkeli  [Philosophy: contradictory observation in relation to the paradigm](#)

**anomalía** (*Kielitiede*) *merkitysten yhteensopimattomuudesta johtuva luonnottomuus*  
Lue koko artikkeli  [Linguistics: unnaturalness due to incongruity of meanings](#)

**anomalía** (*Biologia*) *poikkeavuus tai epämuodostuma*  
Lue koko artikkeli  [Biology: deviancy or malformation](#)

**anomalía** (*Tähtitiede*) *taivaankappaleen rataliikkeeseen liittyvä kulma, joka mitataan radan perihelistä eli pisteestä, jossa aurinkoa kiertävä kappale on lähinnä Aurinkoa*  
Lue koko artikkeli  [Astronomy: the angle between the direction of periapsis and the current position of an object on its orbit](#)

# *P*

**prepositio** (Kielitiede) Lue koko artikkeli



Linguistics: preposition

**postpositio** (Kielitiede) Lue koko artikkeli



Linguistics: postposition

**fosfori** (Biologia) Lue koko artikkeli



Biology: phosphorous

**sähköteho** (Sähkötekniikka) Virtapiirin avulla tai sen eri osissa tapahtuvan energian siirron, syötön ja kulutuksen

*nopeus.  $P = dW/dt = UI$ , missä  $U$  ja  $I$  ovat piirin (osan) napojen välinen jännite ja sähkövirta piirissä (osassa).*

*Sähkötehon SI-yksikkö on watti (W).*

Electrical engineering: electric power

*Tavallisia kerrannaisia ja alikerrannaisia: MW, kW, mW.*

Lue koko artikkeli



# What can be the designation of a concept like?

- What can be the part of speech of a term?
- When can the term be plural?
- Finnish designations (35 089) in the BFT consist of
  - Adjectives 2 377, e.g., [läpinäkyvä](#), transparent
  - Verbs 90, e.g., [haihtua](#), evaporate
  - Adverbs 36, e.g., [alavirtaan](#), downstream
  - Nouns 32 586
    - Property names 782 (2 %) e.g., [biohajoavuus](#), biodegradability  
(Pitkänen-Heikkilä 2015)
    - Action names 1699 (5%) e.g., [metsitys](#), afforestation
  - Plurals 1508 (4,3 %)

# Concept page in the BFT

Kasvitiede **Discussion** Read Edit Edit source **View history** ☆ More Search

## kukkalehti

**kukkalehti**

**Definition** kukan lehti: kehä-, verho-, terä-, emi- tai hedelehti

**Explanation** Kukkalehti on lehtiseuraannon viimeinen lehtityyppi ylälehden jälkeen ja se syntyy kukkaversoon. Kukkalehdet ovat muuttunnaislehtiä, joilla on lisääntymis- ja pölytysbiologinen tehtävä. Ne ovat samasyntyisiä kasvuverson lehtien kanssa. Kehä-, verho- ja terähdet ovat steriilejä lehtiä, hede- ja emilehdet fertilejä siitoslehtiä. Kukan lehdet sijaitsevat toisiaan lähellä kukkapohjukseen kiinnittyneinä ja niiden lehtiasento sekä keskinäiset suhteet vaihtelevat.

### Equivalents


floral leaf englant  
blomblad ruotsi

### Related Concepts


- lehti (yläkäsité)
- lehtiseuraanto
- kehälehti (alakäsité)
- verholehti (alakäsité)
- terälehti (alakäsité)
- emilehti (alakäsité)
- hedelehti (alakäsité)
- kukka (kokonaisuuskäsité)
- kukkapohjus (koostumussuhteinen vieruskäsité)

### Sources

Kalela1971, PyykköM1980, RK1998, BS2001, SLT2006, TeräväE&KanervoE2008



Nurmikaunokilla (*Centaurea phrygia*) on mykerökukinnossaan kahdenlaisia kukkia, suppilomaisia laitakukkia ja torvimaisia kehräkukkia.



Lapinvuokolla (*Dryas octopetala*) on kahdeksan terälehteä.

Categories: Käsitteet Kasvitiede Kasvimorfologia Kukkaan liittyvät käsitteet

Categories: Botany, Plant morphology, Concepts associated flower

hierarkiasuhde | hierarkkinen käsitesuhde | hierarkkinen  
suhde | geneerinen suhde

hierarkiasuhde

hierarkkinen käsitesuhde

hierarkkinen suhde

geneerinen suhde

**Definition** *käsitesuhde*, jossa yhden käsitteen *sisältö* muodostuu toisen käsitteen sisällöstä sekä lisäksi yhdestä tai useammasta erottavasta piirteestä (TSK36)

## Equivalents

generic relation                      englanti

genus-species relation      englanti

|                    |        |
|--------------------|--------|
| relation générique | ranska |
|--------------------|--------|

generisk begreppsrelation ruotsi

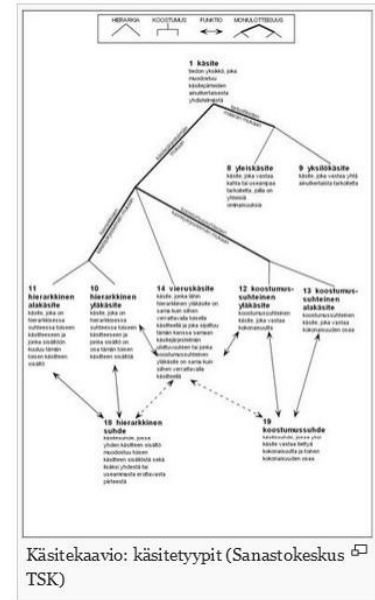
logisk begreppsrelation ruotsi

hierarchische Beziehung saksa

## Related Concepts

- **käsitesuhde** (yläkäsite)
- **koostumusuhde** (vieruskäsite)
- **funktiosuhde** (vieruskäsite)
- **yläkäsite** (funktiosuhde)
- **alakäsite** (funktiosuhde)

## Sources

TSK<sub>36</sub>, TSK<sub>9</sub>



## arkeoni | arkki | arkkieliö

arkeoni (suositeltu) (recommended)

arkki

arkkieliö

arkkibakteeri (vanhentunut) (outdated)

**Definition** arkeonit ovat yksisoluisia, tumattomia eliöitä, joiden solukalvojen rasvoissa on eetterisidokset<sup>[1]</sup>

**Explanation** Arkeonit muodostavat yhden eliökunnan kolmesta pääryhmästä, Archaea (kutsuttiin aiemmin arkeiksi, arkkieliöiksi tai arkkibakteereiksi). Suomenkielinen termi ei ole vielä vakiintunut.

**Scientific Name** *Archaea*

## Equivalents

archaean englantti

archaeon englantti

arké ruotsi

## Related Concepts

- bakteeri (vieruskäsite)
- tumallinen (vieruskäsite)

## Sources

BS2001, SB2013, MBS2014

# Presenting homonymy, polysemy and synonymy

37 630 concept pages, 140 600 designation pages (21.10.2015)

→ **totally 296 440 separate terms**

Homonymy in the BFT: one designation, two or more concepts

- presented on the designation page of a character string

Polysemy inside one namespace: one concept, two or more definitions

- slightly different definitions from a different theoretical direction can be given on the same concept page in the BFT

Synonymy

- designations on a concept page (one concept, two or more designations in the same language)
- if synonymous designations slightly differ, the usage note should be given

*feedback*

**palaute** (Kielitiede) puhujan aivojen puhakeskuksiin palautuva tieto omasta puhetoiminnasta ja äänen perustella tapahtuva puhetoiminnan kontrolli ja ohjaus  
Lue koko artikkeli

**kuittausvuoro** (Kielitiede) opetuskeskustelun kolmasosien vuorottelujärjestyksen kolmas vuoro  
Lue koko artikkeli

**takaisinkytkentä** (Ilmatiede) Lue koko artikkeli

**palaute** (Biologia) Lue koko artikkeli

**takaisinsyöttö** (Biologia) Lue koko artikkeli

**palaute** (Ympäristötieteet) Systemaattisesti esiintyvä lenkki, jossa jokin reaktioketjun osatekijä palauttaa informaatiota tai muuten vaikuttaa ketjun aikaisempiin osiin.  
Lue koko artikkeli

**homonyms**

**huuli**

**Definition**

1. **huulimaisen teriön** tai **verhiön** erillinen ylä- tai alaosa

2. käämmekkäkasvien kukkien huulimainen rakenne, joka muodostuu muuntuneesta kolmannesta lehtisestä osasta

**polysemy**

**Equivalents**

**labellum** englanti

**lip** englanti

**synonyms**

# Definition: What is the most essential thing?

Because of different perspectives, the same concept can be defined differently in relative domains:

|                      |  |
|----------------------|--|
| humus (Botany)       | = soil surface layer formed of partly decomposed organic matter  |
| humus (Soil science) | = organic matter, which is dark and amorphous material formed by high molecular weight biopolymers that are produced during the microbial decomposition and humification of plant and animal residues over a long time |

How should a *WELL* be defined in archaeology?

'a construction where water is kept'



- What does the possible location of a well during excavation look like?
- For what purposes has well been used?
- What material has the well been built of over the course of time?

# What is the most important concept relation from the domain's point of view?

## 1) Generic relation 73 %

**raceme** = an unbranched, elongated **inflorescence** with stalked flowers

important relation between sub- and superordinate concepts  
*raceme : inflorescence*



Fin. terttu

## 2) Partitive relation 26 %

**flower** = a structure of a **plant** that contains **floral leaves**

important relation between comprehensive and partitive concepts  
*flower : floral leaf*



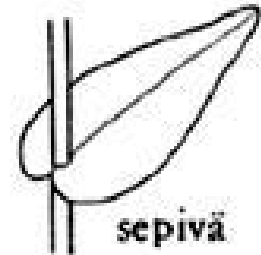
# What is the most important concept relation from the domain's point of view?

Associative relations, e.g.

3) Property relation 20 %

**amplexicaul** = (of a **leaf**) clasping the stem

important relation between property and property's owner  
*amplexicaul: leaf*



4) Birth relation 4 %

**fruit** = a part of a **plant** developed from an **ovary**

important relation between result and originator  
*fruit : ovary*



# Can we always find the superordinate concept?

sign (semiotics) = something that represents something else

- Theoretical concepts in humanities are often defined without a superordinate concept in BFT
  - head word of the definition from the general language
  - definition often includes other terms; often the concept relation is associative
- Experiment in forming an ontology of plant morphology terms
  - converting terminological material to an ontology is not simple if the concept relations are to be presented as exactly as specialists have presented them in the BFT's wiki
  - BFT's material shows that botanical concepts do not always have a generic superordinate concept or they can have many superordinate concepts

# Definition of concepts vs. term type

Does the form of designation have an effect on the definition?

- Complex terms (in Finnish, total of 17 623 [50%] in the BFT)
  - Some difficulties, unusually many hidden definitions
- Simple terms (9325; 25%)
- Compound terms (two or more separate words in Finnish, 7773 [21%])
  - Some difficulties, unusually many hidden definitions
  - Long complex phrases are not clearly differentiated from the definition
- Abbreviations (1388; 4%)
  - Experts do not remember to define a concept, "an acronym of X"
- Also some direct references to the designations occur in the BFT  
"designation/general designation of X" or "general concept for X"

# Explanation

- Space is unlimited: examples, quotations, concept history can be given
- Place for theoretical speculation and discussion
- Possibility of describing differences between special and general language

Explanation:

Gymnosperms do not have flowers although they can appear to flower.



## kukka

kukka

**Definition** kukkapohjuksen ja kukkalehtien muodostama siemenkasvien rakenne

**Explanation** Verson osa, jonka tehtävä on huolehtia kasvin lisääntymisestä.

Täsmällisesti rajattuna *kukka* tarkoittaa vain **koppisiemenisten** kasvien lisääntymiseen erilaistunutta rakennetta, vaikka arkikielessä usein puhutaan myös **paljassiemienisten** kukasta ja kukkimisesta.

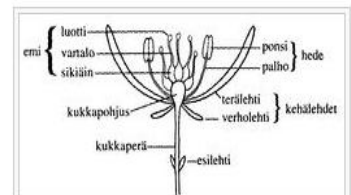
Paljassiemienisillä ei ole kukkaa, mutta niilläkin voidaan ajatella esiintyvän kukintaa (esim. **hedekäpy**, **emikäpy**, siemenkäpy, hedesuomu, emisuomu, siemensuomu).

## Equivalents

|               |  |
|---------------|--|
| flower        | englanti   |
| floret        | englanti (tiividen kukintojen yksittäisistä kukista, suppeampi käsite) |
| blóma         | fääri  |
| blóm          | islanti  |
| smávaxið blóm | islanti  |



Järvisätkin *Ranunculus peltatus*



Retkeilykasvio 1998



# Equivalents: terms in other languages

- Possibility of offering equivalents in different languages unlimited
  - Equivalents in 35 different languages, e.g., Inari Sami, Finnish Romany, Faroese, Japanese
- Possibility of giving information about equivalency

Finnish: kukka

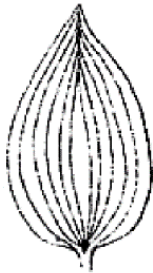
| Equivalents |   |
|-------------|---|
| flower      | englanti  |
| floret      | englanti (tiiviiden kukintojen yksittäisistä kukista, suppeampi käsite) |

- There may be differences in term equivalents
  - native/indigenous in biology is dependent on the area
  - linguistic concepts are dependent on language



# A picture is worth a thousand words

*"In a multimodal conceptual description, we believe that the structured information in terminographic definitions should mesh with the visual information in images for a better understanding of complex and dynamic concept systems." (Faber et al. 2007.)*



ovate = with a curved outline and widest below the middle  
= 1.2-2 times as long as wide

Citizen science: wide crowdsourcing possible

# Naming of concept relations

- A field for linking Related Concepts and specifying the concept relation
- Naming of concept relations seems to be difficult for the experts of the BFT
- A concept system built by experts could later help in mapping gaps from the data

|   |   |
|---|---|
| <b>sieni</b>  |   |
| sieni   | fungus  |
| <p><b>Definition</b> tumallinen, lehtivihreätön eliö, joka hajottaa kuollutta eloperäistä materiaalia tai elää loisena</p> <p><b>Explanation</b> Sienet ovat yksi- tai monisoluisia ja jälkimmäisessä tapauksessa perusrakenteeltaan rihmamaisia. Rakenteita ympäröi solunseinä, joka sisältää kitiiniä tai selluloosaa tai molempia yhdessä monien muiden monimutkaisten orgaanisten molekyylien kanssa. Sienirihmoista eli hyyfeistä kehittyy monissa ryhmissä ajoittain pseudoparenkymistä muodostuvia itiöemiä, joiden yhteydessä syntyy itiöitä (sieni-itiöitä). Itiöiden lisäksi sienet lisääntyvät suvuttomien, rihmoista kuroutuvien kuromaitiöiden eli konidioiden välityksellä.</p> |   |
| <b>Equivalents</b>  |   |
| fungus latina   |   |
| <b>Related Concepts</b>   |   |
| <ul style="list-style-type: none"> <li>• eliö (yläkäsité)</li> <li>• kasvi (vieruskäsité)</li> <li>• eläin (vieruskäsité)</li> <li>• mikrobi (vieruskäsité)</li> </ul>  | <ul style="list-style-type: none"> <li>organism (superordinate c.)</li> <li>plant (coordinate c.)</li> <li>animal (coordinate c.)</li> <li>microbe (coordinate c.)</li> </ul> |

# Supporting crowdsourced volunteers and all users

- The BFT has
  - 300 experts in various namespaces
  - 300 other registered users
  - 1500—2000 individual users on every working day (Google Analytics)
- Maintenance and organization needed: support for terminology work, training for experts, university courses for doctoral students
- Automatization: every participant is building one piece of the large whole, the machine should connect each new piece to the right place in the process

# Target for development: Help for terminology work of experts

Can something be automatized?

- searching for related concepts; help with creating a concept map
- offering equivalents (e.g. from E-Thesis database)
- searching for context examples from other database; linkage

Notwithstanding paragraph 7 above, a template or other **representation** of a biometric of the holder which meets the specification provided for in Annex I shall also be required for  
ment and it shall not be possible to reconstitute it from the template or other **representation**;

At the same time, low-quality housing and over **representation** of immigrants in deprived urban neighbourhoods create problems for integration in most Member States.

A balanced gender **representation** should be promoted.

Improve **representation** and manage co-ordination of EU-positions in international bodies | European Commission | ongoing |

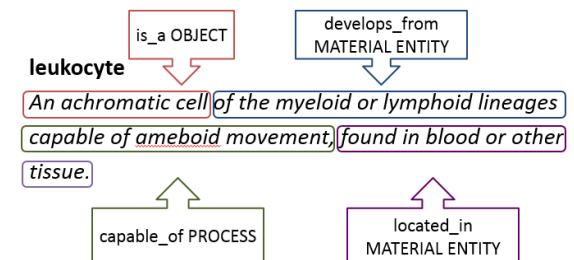
nd consultations Other expenditure ( specify ) | 0.050 0.090 0.100 0.010, frais de **representation** |

In their appointment account shall also be taken of diversity of geographic **representation** and of legal expertise.

countries and the long-term trend is well that of an improvement of women's **representation** in politics.

protection, notably through improved evidence, better consultation and better **representation** of consumers ' interests.

- creating tools that help in writing definitions  
(e.g., Sager & L'Homme 1994, Seppälä 2015)



# Pros and cons of the wiki platform

- + flexible
- + builds real hypertext
- + easily found by search engines
- + working is transparent
- always under construction
- shared concept can be edited only in the namespace where it is sited
- entry pages have to be named

The following 1,000 pages are in this category, out of 5,166 total.

([previous page](#)) ([next page](#))

## E

- [Biologia:elävät fossiilit](#)
- [Biologia:EM](#)
- [Biologia:emaskulointi](#)
- [Biologia:Embden-Meyerhofin reaktiotie](#)
- [Biologia:embryoblasti](#)
- [Biologia:embryoidi](#)
- [Biologia:embryologia](#)
- [Biologia:embryonaalinen](#)
- [Mikrobiologia:fotoheterotrofi](#)
- [Biologia:fotometria](#)
- [Biologia:fotomorfogeneesi](#)
- [Biologia:fotonastia](#)
- [Fysiikka:fotoni](#)
- [Biologia:fototaksia](#)
- [Biologia:fototropismi](#)
- [Biologia:fouling-organismi](#)
- [Biologia:fovea](#)
- [Biologia:fraktaali](#)
- [Biologia:harppauskerros](#)
- [Biologia:harvinainen](#)
- [Biologia:haulikkoammuntakoe](#)
- [Biologia:hedelmällisyys](#)
- [Biologia:hedelmöitys](#)
- [Biologia:heidelberginihminen](#)
- [Biologia:heikkenevän luonnonvalinnan teoria](#)
- [Biologia:heimo](#)
- [Biologia:heksaani](#)
- [Biologia:heksaanihappo](#)

# Research project: Quality of terminology work built by experts' crowdsourcing

- Are term entries good?
  - Have the boundaries of concepts been presented clearly?
  - Are the equivalents of the terms correct and are the differences between terms in different languages described?
  - How systematic are definitions and how clear are explanations?
  - Are the concept relations presented correct?
  - Questionnaires to expert users and ordinary users
- How can terminological information be used?
  - In teaching
  - In research
  - In various digital applications
  - Questionnaire to users having different user profiles

# References

ISO 704: 2009. Terminology work - Principles and methods.

Faber, Pamela, Pilar León Araúz, Juan Antonio Prieto Velasco, and Arianne Reimerink. 2007. Linking Images and Words: the description of specialized concepts. *International Journal of Lexicography* 20, no. 1: 39-65.

Pitkänen-Heikkilä, Kaarina 2015. Adjectives as terms. *Terminology* 21, 1. 76-101.

Sager, Juan & Marie-Claude L'Homme. 1994. A model for the definition of concepts: Rules for analytical definitions in terminological databases. *Terminology*, 1(2):351–373.

Seppälä, Selja 2015: An ontological framework for modeling the contents of definitions. *Terminology* 21, 1. 23-50.

Terminologian sanasto 2006. Finnish Terminology Centre TSK.

Virta, Tuuli 2015: Tieteen kansallinen termipankki ja termien määritelmät. Unpublished master's thesis. University of Helsinki.